

Investigating the Role of Histidine 157 in the Catalytic Activity of Human Cytomegalovirus Protease[†]

Reza Khayat,[‡] Renu Batra,[‡] Marie-Josée Massariol,[§] Lisette Lagacé,[§] and Liang Tong^{*,‡}

Department of Biological Sciences, Columbia University, New York, New York 10027, and Bio-Méga Research Division, Boehringer Ingelheim (Canada) Ltd., 2100 Cunard Street, Laval, Quebec H7S 2G5, Canada

Received January 24, 2001; Revised Manuscript Received March 21, 2001

ABSTRACT: Herpesvirus proteases belong to a new class of serine proteases and contain a novel Ser-His-His catalytic triad, while classical serine proteases have an acidic residue as the third member. To gain a better understanding of the molecular basis for the functional role of the third-member His residue, we have carried out structural and biochemical investigations of human cytomegalovirus (HCMV) protease that bears mutations of the His157 third member. Kinetic studies showed that all the mutants have reduced catalytic activity. Structural studies revealed that a solvent molecule is hydrogen-bonded to the His63 second member and Ser134 in the H157A mutant, partly rescuing the activity of this mutant. This is confirmed by our kinetic and structural observations on the S134A/H157A double mutant, which showed further reductions in the catalytic activity. The structure of the H157A mutant is also in complex with the PMSF inhibitor. The H157E mutant has the best catalytic activity among the mutants; its structure, however, showed conformational readjustments of the His63 and Ser132 residues. The Ser132-His63 diad of HCMV protease has similar activity as the diads in classical serine proteases, whereas the contribution of the His157 third member to the catalysis is much smaller. Finally, structural comparisons revealed the presence of two conserved structural water molecules at the bottom of the S₁ pocket, suggesting a possible new direction for the design of HCMV protease inhibitors.

Herpesviruses are widely distributed in nature and afflict many species throughout the animal kingdom (1). Eight herpesviruses are currently known to infect humans, of which the human cytomegalovirus (HCMV)¹ can cause severe health problems in immuno-suppressed and immuno-compromised individuals. The mature nucleocapsids of herpesviruses contain the viral genome as double-stranded DNA and are enveloped by a membrane bilayer that also contains virally encoded glycoproteins. The formation of this nucleocapsid requires a scaffold that is produced by the virally encoded assembly protein. An intermediate capsid is then constructed around this assembly protein scaffold. This is followed by a critical step in the viral life cycle—proteolytic processing of the assembly protein and the encapsidation of the viral DNA (2, 3). Herpesvirus mutants that have a defect in this proteolytic processing can only form aberrant empty capsids, devoid of the viral DNA.

Genetic and biochemical studies showed that the proteolytic processing of the assembly protein is catalyzed by

a protease that is encoded within the viral genome (2). The herpesvirus protease is a serine protease, although its sequence does not bear any homology to other serine proteases or other proteins in general. The protease has a molecular mass of about 30 kDa, with 256 amino acid residues for HCMV protease, and prefers small residues at the P₁ (Ala) and P₁' (Ser) positions. The protease exists in a monomer–dimer equilibrium in solution, and the dimer is the active species of the enzyme (3). The activity of the protease can be enhanced significantly by kosmotropic agents such as glycerol and Na₂SO₄, due in part to the stabilization of the dimer species of the enzyme (3).

Structural studies of the herpesvirus protease-free enzyme showed that it belongs to a new class of serine proteases, consistent with its unique amino acid sequence and biochemical properties (4–10). The backbone fold of the protease consists of a central, mostly antiparallel, seven-stranded β -barrel (β 1 through β 7) that is surrounded by seven helices (α A through α G), distinct from the structures of classical serine proteases (such as chymotrypsin and subtilisin). The catalytic nucleophile, Ser132 (residue numbering in this paper is according to HCMV protease), is located on the surface of the β -barrel. In contrast to classical serine proteases, the structures revealed that herpesvirus proteases possess a novel catalytic triad, Ser132-His63-His157, with a histidine residue (His157) as the third member. An Asp or Glu residue is normally found as the third member in the catalytic triad in classical serine proteases. The region near the active site in the free enzyme structure is rather featureless, with little indications for the locations of the

[†] This work was supported by the National Institutes of Health (Grant AI46139 to L.T.). R.K. was supported by an NIH training program in molecular biophysics (GM08281).

* Correspondence should be addressed to this author. Phone: (212) 854-5203, FAX: (212) 854-5207, E-mail: tong@como.bio.columbia.edu.

[‡] Columbia University.

[§] Boehringer Ingelheim (Canada) Ltd.

¹ Abbreviations: APS, Advanced Photon Source; CHESS, Cornell High Energy Synchrotron Source; E, enzyme; HCMV, human cytomegalovirus; MOPS, 3-morpholinopropanesulfonic acid; NSLS, National Synchrotron Light Source; PMSF, phenylmethylsulfonyl fluoride; rms, root-mean-square; S, substrate.

Table 1: Summary of Crystallographic Information

structure	H157E	H157Q	H157A	S134A	H157A/S134A
X-ray source	APS ComCAT	NSLS X4A	CHESS A-1	APS ComCAT	NSLS X4A
max resolution (Å)	2.0	2.2	2.2	2.3	2.0
no. of observations	136174	66653	88643	108421	94409
R_{merge} (%) ^a	9.7	4.9	7.7	7.0	8.8
cell parameters (<i>a</i> , <i>c</i>) (Å)	75.9, 170.6	76.1, 169.4	76.2, 170.6	76.4, 171.8	76.0, 167.6
resolution range	20–2.0	20–2.2	20–2.2	20–2.3	20–2.0
no. of reflections	30313	25469	25276	22695	32752
completeness (%) (2σ cutoff)	87	98	97	97	97
R /free R factor ^b (%)	22.7/26.9	22.0/26.0	22.6/26.9	22.9/27.0	22.6/25.9
rms deviation in bond lengths (Å)	0.006	0.006	0.006	0.006	0.005
rms deviation in bond angles (deg)	1.2	1.2	1.2	1.2	1.2

$$^a R_{\text{merge}} = \frac{\sum_h \sum_i |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i I_{hi}}. \quad ^b R = \frac{\sum_h |F_h^o - F_h^c|}{\sum_h F_h^o}$$

substrate binding pockets (4). The structure of HCMV protease in complex with a peptidomimetic inhibitor revealed that the protease undergoes large conformational changes upon inhibitor binding (11). This is consistent with solution studies that showed a blue shift in the tryptophan fluorescence emission spectra upon inhibitor binding (12). Therefore, the herpesvirus protease appears to have an induced-fit behavior, in contrast to classical serine proteases where substrate binding generally has little impact on the conformation of the enzyme (13). Nonetheless, the inhibitor is bound to HCMV protease by forming an antiparallel β -sheet with the enzyme, and this binding mode is similar to that observed for classical serine proteases (11).

The third member of the catalytic triad in classical serine proteases is critical for the enzymatic activity. For example, mutation of the Asp102 residue of trypsin to Asn led to a 38 000-fold reduction in the k_{cat} of the enzyme, while maintaining the K_m value (14). Similarly, mutation of the third-member Asp residue to Ala in subtilisin reduced the k_{cat} by 20 000-fold (15). In contrast, biochemical data suggest that the role of His157 in the catalysis by herpesvirus proteases may be different from the classical serine proteases. First of all, a His157 to Ala mutation caused only a limited decrease in the catalytic efficiency of the protease (16). Second, replacing the His157 residue by Glu or Asp, in an effort to mimic the triad in classical serine proteases, did not enhance the activity of the enzyme (4). To obtain a better understanding of the functional role of the His157 residue in the catalysis by herpesvirus proteases, we have characterized the His157 mutants of HCMV protease in greater structural and biochemical detail.

MATERIALS AND METHODS

Preparation of Enzymes. The protease samples used in these studies all contain the Ala143Gln (A143Q) mutation, which eliminates the autoproteolysis at this position (17). This mutant essentially maintains the catalytic activity of the native form. Protease samples carrying the A143Q change will be referred to as the wild type here. The protease was expressed in *E. coli* strain BL21(DE3) pLysS, and the purification followed the protocol that was described earlier (4). Site-directed mutagenesis was performed with the QuikChange kit (Stratagene), and the presence of the mutations was confirmed by sequencing. The mutants were expressed and purified following the same protocol as that for the A143Q sample.

Crystallization, Data Collection, and Data Processing. Crystals were grown at 21 °C using the hanging drop vapor diffusion method from a solution containing 1 μ L of 14 mg/mL protein [in 20 mM NaOAc (pH 5.0), 80 mM NaCl] and 1 μ L of the reservoir solution [19–21% PEG 3350, 0.1 M MES (pH 6.0), 15% glycerol, 5% *tert*-butyl alcohol, and 0.3–0.4 M NaCl]. The crystals belong to the space group $P4_12_12$ with unit cell dimensions of $a = b = 76$ Å and $c = 170$ Å (Table 1). Crystals were flash-frozen in liquid propane for data collection at 100 K. X-ray diffraction data up to 2 Å resolution were collected at synchrotron radiation sources, including Advanced Photon Source (APS) beam line 32-ID (ComCAT), National Synchrotron Light Source (NSLS) beam line X4A, and Cornell High Energy Synchrotron Source (CHESS) beam line A-1 (Table 1). The diffraction images were processed with the HKL package (Table 1) (18).

Structure Determination and Refinement. The structure of the H157A mutant was determined by molecular replacement with the REPLACE program suite (19) and the free enzyme dimer as the search model (4). Residues involved in the packing of the free enzyme crystal (23–33) were removed. The structure model was then subjected to rigid-body refinement using reflections between 5.0 and 3.5 Å resolution. Structure factors were calculated for all reflections to 2.5 Å resolution using the atomic model after rigid-body refinement, and the calculated phases were applied to the observed structure factor amplitudes. The phase information was improved by 2-fold noncrystallographic symmetry-averaging over the dimer with the program DM (20). The atomic model was rebuilt against the resulting electron density map with the program O (21). It clearly showed the positions of residues 23–33 in both monomers, and revealed that the catalytic Ser132 residue had been covalently modified. During the purification of this batch of samples, the H157A mutant was incubated overnight with 2 mM phenylmethylsulfonyl fluoride (PMSF). Therefore, it is likely that Ser132 has been modified by PMSF, and the observed electron density is generally consistent with a PMSF molecule. This batch of protein samples was used only for crystallization and structural studies, and not for kinetic studies. The structure refinement of the H157A mutant was carried out with the program CNS (22). The structure refinement of the other mutants was initiated with the structure of the H157A mutant. The crystallographic information is summarized in Table 1.

Kinetic Measurements. The k_{cat} and K_m values for each enzyme were measured by monitoring the hydrolysis of the

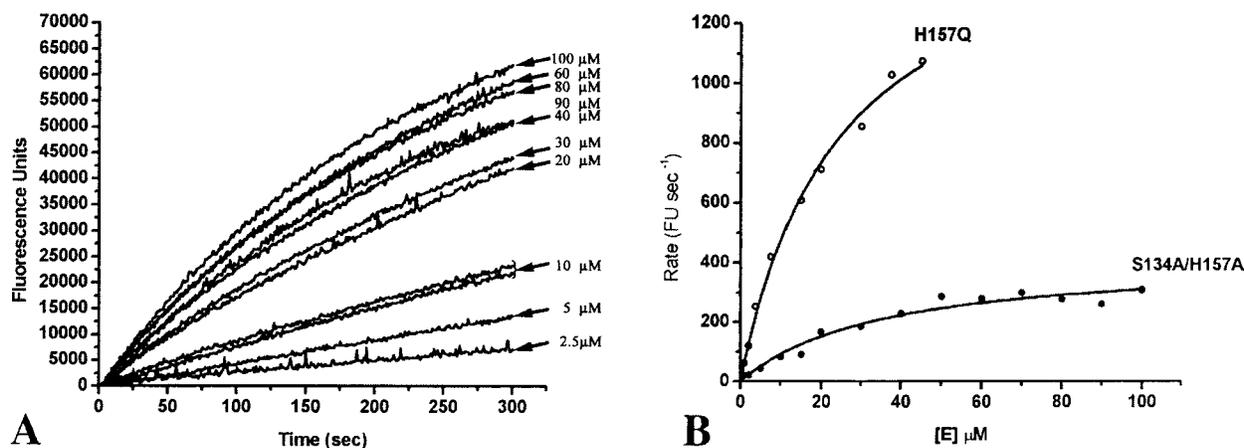


FIGURE 1: Characterization of HCMV protease mutants using a new kinetic protocol. (A) Reaction progress curves for the S134A/H157A double mutant. The substrate concentration was fixed at $0.25 \mu\text{M}$, and the enzyme concentration was varied between 2.5 and $100 \mu\text{M}$. (B) Nonlinear least-squares fit to the kinetic equation for the H157Q and S134A/H157A mutants. The initial velocity is plotted as a function of the total enzyme concentration. The k_{cat} and K_{m} parameters are obtained directly from this fitting (program Origin, Microcal Software, Inc.).

Table 2: Summary of Kinetic Parameters

enzyme	k_{cat} (s^{-1}) ^a	K_{m} (μM) ^a	$k_{\text{cat}}/K_{\text{m}}$ ($\text{s}^{-1} \text{M}^{-1}$) ^a
wild-type (with A143Q)	0.033	10.1	3,300
H157E	0.0041 (8)	14.5 (0.7)	283 (12)
H157Q	0.0036 (9)	20.3 (0.5)	177 (19)
H157A	0.0033 (10)	21.8 (0.5)	151 (22)
S134A	0.00061 (54)	27.7 (0.4)	22 (150)
H157A, S134A	0.00069 (48)	35.6 (0.3)	19 (170)
trypsin (D102N)	0.0011 (38000)	90 (1.6)	12 (24000)
subtilisin (D32A)	0.0023 (19000)	480 (0.4)	4.8 (52000)

^a The numbers in parentheses are the ratios between the wild-type and mutant values. See Materials and Methods for details on the kinetic assay for HCMV protease. Values for the trypsin and subtilisin mutants are from the literature (14, 15).

internally quenched peptide substrate 4-(4'-dimethylaminophenylazo)benzoyl-Arg-Gly-Val-Val-Asn-Ala-↓-Ser-Ser-Arg-Leu-Ala-(2'-aminoethyl)amino-naphthalene-1-sulfonic acid (Bachem) at 30°C ($\lambda_{\text{ex}} = 355 \text{ nm}$ and $\lambda_{\text{em}} = 495 \text{ nm}$) (23). The reaction mixture was composed of 100 mM MOPS (pH 7.2), 0.5 M Na_2SO_4 , 10% glycerol, and 10 mM DTT. The fluorescence of the cleaved peptide was monitored using a Photon Technologies Inc. spectrofluorometer. The kinetic parameters for the wild-type enzyme were obtained from classical Michaelis–Menten kinetics. The mutants studied here, however, have significantly reduced catalytic activity, requiring high enzyme concentrations (up to $3 \mu\text{M}$) for reliable measurement of the reaction rate. This, however, violates the basic assumption of the kinetic analysis ($[\text{S}] \gg [\text{E}]$), as the substrate concentration needs to be maintained below $10 \mu\text{M}$ for solubility and to avoid quenching by substrates. Therefore, a different protocol was used to obtain the kinetic parameters of these ‘slow’ mutants (R.B., R.K., and L.T., unpublished data). In these assays, the substrate concentration was held constant (at $0.25 \mu\text{M}$), and the enzyme concentration was varied between 2.5 and $100 \mu\text{M}$ ($[\text{E}] \gg [\text{S}]$, and with $[\text{E}]$ varying around the expected K_{m}). The kinetic parameters can then be extracted from such kinetic studies (see below), using nonlinear least-squares fitting (program Origin, Microcal Software, Inc.) (Figure 1, Table 2).

RESULTS AND DISCUSSION

Kinetic Studies of Wild-Type (A143Q) and H157 Mutants. The kinetic properties of the wild-type (A143Q) and H157 mutants of the protease were characterized using a fluorogenic substrate (23). A new experimental protocol was used to obtain the kinetic parameters for the mutants as they have significantly reduced catalytic activity (R.B., R.K., and L.T., unpublished data). This protocol is based on the fact that the Michaelis–Menten analysis is symmetrical with respect to the substrate (S) or the enzyme (E). Therefore, the reaction velocity as a function of $[\text{E}]$ is given by

$$v = k_2[\text{S}]_0 \left(1 + \frac{K_{\text{m}}}{[\text{E}]} \right)$$

This is the same form as that of the classical Michaelis–Menten equation, as a function of $[\text{S}]$. Therefore, with this new protocol, $[\text{E}]$ is varied around the expected K_{m} value, and in large excess over $[\text{S}]$. The analysis of v as a function of $[\text{E}]_0$ will then produce the k_{cat} and K_{m} parameters (Figure 1, Table 2). This analysis can use exactly the same mathematical equations as the traditional kinetic protocol, such as Lineweaver–Burk plots or preferably nonlinear regression. The only requirement of this new protocol is that there cannot be significant substrate turnover during the reaction, such that $[\text{S}]_0$ can be treated as a constant. Therefore, this protocol is ideally suited to the kinetic studies of ‘slow’ enzymes and especially active site mutants, as many of these will have severely reduced catalytic activities. In addition, the reaction rate used to fit the equation above should be the initial velocity, also to minimize substrate turnover.

All the H157 mutants have a reduction in the catalytic activity. Detailed analysis showed that the reduction in activity is due mostly to a drop in the k_{cat} of the enzyme, whereas the K_{m} remains essentially unchanged (Figure 1, Table 2). This is consistent with observations from the structure of the inhibitor complex that the His157 side chain is not directly involved in the binding of peptidomimetic inhibitors (11). Therefore, mutation of the His157 residue is likely to affect only the catalysis itself, through its involve-

ment in the catalytic triad, and is unlikely to affect the binding of substrates. The H157A mutant shows the largest change, with a 22-fold reduction in the catalytic activity. The H157E mutant, which would mimic the catalytic triad in classical serine proteases, is more active than the H157A mutant, but still has a 12-fold loss of activity as compared to the wild-type enzyme. It also appears that the negative charge on the Glu side chain is needed, as the H157Q mutant behaves more similar to the H157A mutant in terms of the catalytic activity (Table 2).

In contrast, knocking out the third member of the catalytic triad in classical serine proteases leads to greater than 20 000-fold reduction in the catalytic activity of the enzyme (Table 2) (14, 15). This comparison suggests that the contribution of the His157 third member to the catalysis by herpesvirus protease is much smaller relative to that of the classical serine proteases. This weaker third member in the herpesvirus proteases may be the basis for their slower catalysis. The k_{cat} value for HCMV protease is less than 0.1 s^{-1} with the best peptide substrate that is currently known (24), whereas the k_{cat} values for classical serine proteases are about 60 s^{-1} (14, 15). It may be possible that herpesvirus proteases can become faster with other (peptide) substrates yet to be identified. Even with the natural substrate (the assembly protein precursor), the k_{cat} value is only about 0.2 s^{-1} (17). However, herpesviruses harboring the H157A mutation in the protease can still make mature virions (16), suggesting that the virus may not have a stringent requirement on the catalytic efficiency of the protease.

Interestingly, the k_{cat} values of all the enzymes that carry an inactivating mutation of the third member are similar, in the range of 0.002 s^{-1} for trypsin, subtilisin, and HCMV protease (Table 2). This suggests that the efficiency of the catalytic diad (Ser132-His63) in herpesvirus proteases is similar to that in the classical serine proteases, and gives further evidence that the third member in herpesvirus proteases has a smaller contribution to the catalytic efficiency.

Crystal Structures of the H157 Mutants of HCMV Protease. The crystal structures of the H157E, H157Q, and H157A mutants of HCMV protease have been determined at up to 2 Å resolution (Table 1). The atomic models are of excellent quality, with low crystallographic *R* factors and low deviations from ideal bond length and bond angle parameters (Table 1). A total of 87% of the residues are in the favored regions of the Ramachandran plot, and 13% are in the generously allowed regions. No residues are in the disallowed regions of the Ramachandran plot (data not shown). The atomic coordinates have been deposited at the Protein Data Bank (entries 1ID4, 1IED, 1IEC, 1IEF, and 1IEG).

There is a dimer of HCMV protease in the crystallographic asymmetric unit (Figure 2). The current atomic models contain residues 9–45, 55–137, 153–200, and 211–256 in one monomer of the dimer, and residues 304–345, 355–435, 453–501, 504–508, and 511–556 in the second monomer (Figure 2A). To distinguish between residues in the two monomers, residue numbers for the second monomer have been incremented by 300. As was observed earlier with the free enzyme and the inhibitor complex of HCMV protease (4, 11), several loops of the protease (loops L1, L3, L9) are highly disordered and their conformations could not be determined in the current structures either. However, the

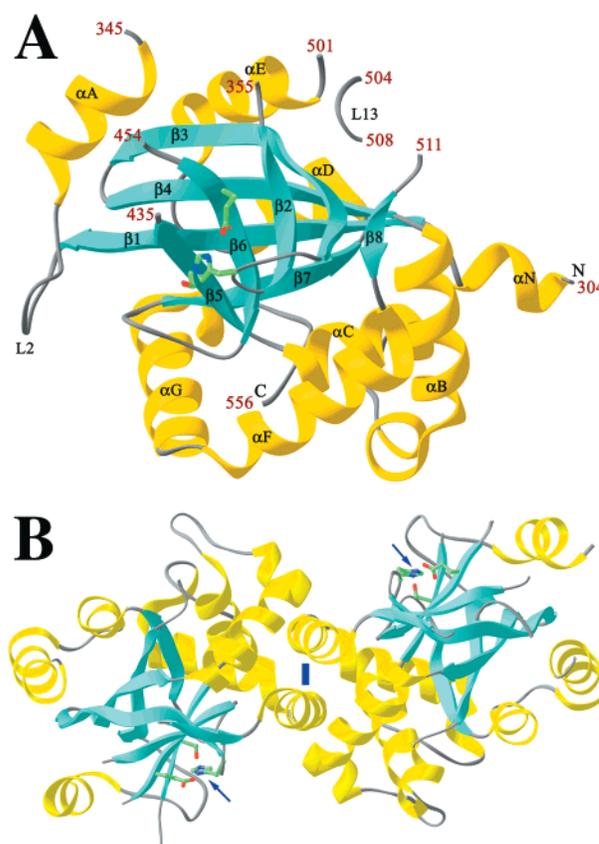


FIGURE 2: Crystal structure of the H157E mutant. (A) Structure of the second monomer of the H157E mutant. The active site residues (Ser432, His363, Glu457) are shown as stick models (in green for carbon atoms). The regions that show large conformational differences to the structures of the wild-type free enzyme are labeled. These include helix αN and loops L2 and L13. The conformations of helix αN and loop L2 are similar to those in the inhibitor complex of HCMV protease (11). (B) Structure of the dimer of HCMV protease. The active sites of the two monomers are indicated by the arrows. The 2-fold axis of the dimer is shown as a blue bar. Both panels produced with Ribbons (30).

conformation of residues 504–508, in loop L13 of the second monomer, was determined in the current structure (Figure 2A). This is the first time residues in loop L13 have been located in any of the HCMV protease structures.

In addition, residues in loop L2 are well ordered and have the same conformation in both monomers in the crystals studied here. The conformation of this loop actually resembles that found in the inhibitor complex of HCMV protease (11). Inhibitor binding appears to stabilize the conformations of residues in loops L2 and L9 (11). In the current structures, however, residues in loop L9 are still disordered. Attempts were made to soak various inhibitors into the current crystals, expecting that loop L9 may become ordered upon inhibitor binding. However, analyses of the resulting X-ray diffraction data and electron density maps showed that the inhibitors did not bind. These inhibitors all contain an activated carbonyl and require the attack by the catalytic Ser132 for inhibition (25, 26). The current crystals are grown at pH 6, which limits the activity of the enzyme and the inhibitory potency of the compounds.

The overall structures of the H157E, H157Q, and H157A mutants are rather similar. The root-mean-square (rms) distance among 430 equivalent C α atoms of any pair of the 3 dimer structures is about 0.3 Å. In each structure, the two

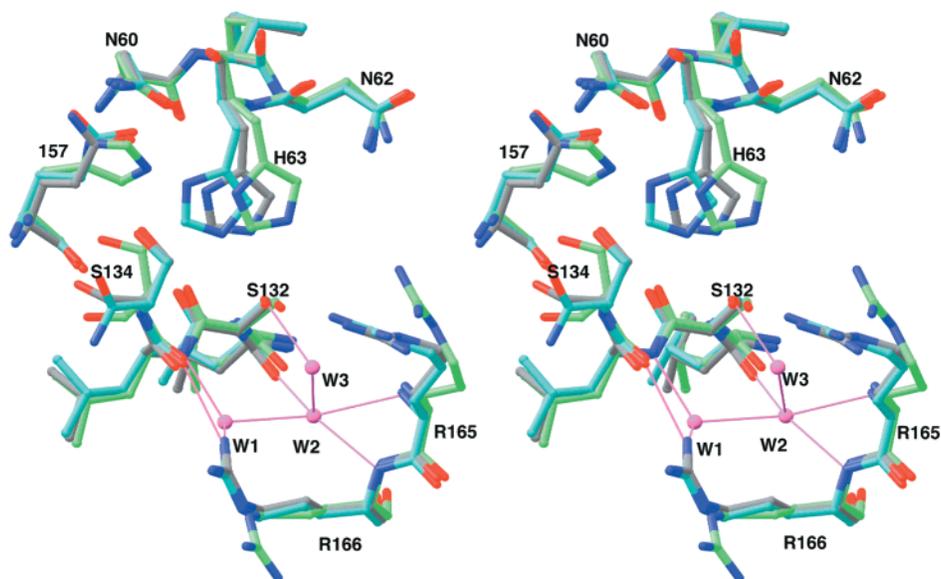


FIGURE 3: Active site of the H157E and H157Q mutants. Stereo drawing showing the overlap of the structures of the wild-type free enzyme (in green for carbon atoms), H157E (in cyan), and H157Q (in gray) mutants, near the active site. Also shown are three conserved water molecules (labeled W1, W2, and W3) near the active site region (purple spheres), and their potential hydrogen-bonding interactions. Produced with Ribbons (30).

monomers of the dimer are related by a rotation of 179.4° , almost an exact 2-fold axis. However, residues 504–508 appear to be ordered only in the second monomer, and residues 304–308, at the N-terminus of the second monomer, form the helix α_N that was also observed in the inhibitor complex structure (Figure 2A) (11). In addition, residues 248–251, in loop L16, assume different conformations in the two monomers, possibly due to crystal packing influences. These differences between the two monomers are located far from the active site of the enzyme, and are unlikely to affect the catalysis by the enzyme.

The organizations of the dimers of the H157 mutants are slightly different than those of the free-enzyme and inhibitor complex that we have reported earlier (4, 11). When one monomer of the wild-type and H157 mutant dimers is superimposed, a rotation of about 3° is needed to bring the other monomer into superposition. With the dimer of the inhibitor complex, a rotation of 3° , but in the opposite direction, is needed to bring the second monomer into superposition. This suggests that the HCMV protease dimer shows a certain degree of variability (4). A much larger change in the organization of the dimer is seen in the structures of the proteases from different subfamilies of herpesviruses (8–10).

Catalytic Triads of the H157E and H157Q Mutants. Within the monomer of the protease, the overall structures of the H157E and H157Q mutants are similar to the wild-type free-enzyme structure (4). The rms distance is 0.4 \AA between 187 equivalent $C\alpha$ atoms. Most of the differences observed between the mutant and the wild-type structures involve residues located on the surface of the enzyme. These differences are also apparent among the structures of the wild-type enzymes (4–7), and therefore may be due to flexibility in the protease and the differences in the crystallization conditions. The reduction in catalytic activity of the H157E and H157Q mutants is therefore unlikely due to an overall change in the enzyme structure as a result of the mutations.

Near the active site region, there are discernible differences in the positioning of the catalytic triad residues, in addition to the mutation at the His157 position (Figure 3). In the first monomer, the conformation of the side chain of Glu157 corresponds to one of its favored rotamers. However, the side chain carboxylate OE1 oxygen atom does not superimpose with the NE2 atom of the His157 side chain. The two atoms are separated by a distance of 1.2 \AA . The NE2 atom of His157 makes the hydrogen bond to the His63 second member in the wild-type protease. In the H157E mutant, the second member His63 residue shows conformational changes in both the main chain and side chain atoms, possibly to maintain the hydrogen bond with the new Glu157 side chain (Figure 3). In its new position, the side chain of His63 in the H157E mutant is in steric clash (within 1.4 \AA) with the side chain of His157 in the wild-type structure (Figure 3). There is also a movement of the side chain of Asn62, which appears to be coupled to that of His63. The side chain hydroxyl group of the catalytic Ser132 residue moves by about 0.6 \AA in the H157E mutant structure to maintain the hydrogen bond with the His63 residue in its new position (Figure 3). In the second monomer of the dimer, similar changes in the conformation of the Ser432 and His363 residues are observed. However, the Glu457 side chain adopts a different rotamer, even though it can still form the hydrogen bond with the His363 side chain.

The largest difference between the structures of the H157Q and H157E mutants is in the conformation of the His63 residue, which is more like the wild-type structure in the H157Q mutant (Figure 3). The conformation of the Gln157 side chain is similar to that of the Glu157 side chain in the first monomer. Some flexibility of the Gln457 side chain amide group was observed in the second monomer of the dimer, where the χ_3 torsion angle could not be precisely defined based on the electron density. Kinetically, the H157Q mutant behaves more like the H157A mutant, possibly due to the weaker hydrogen-bonding interactions between His63 and Gln157 as compared to Glu157.

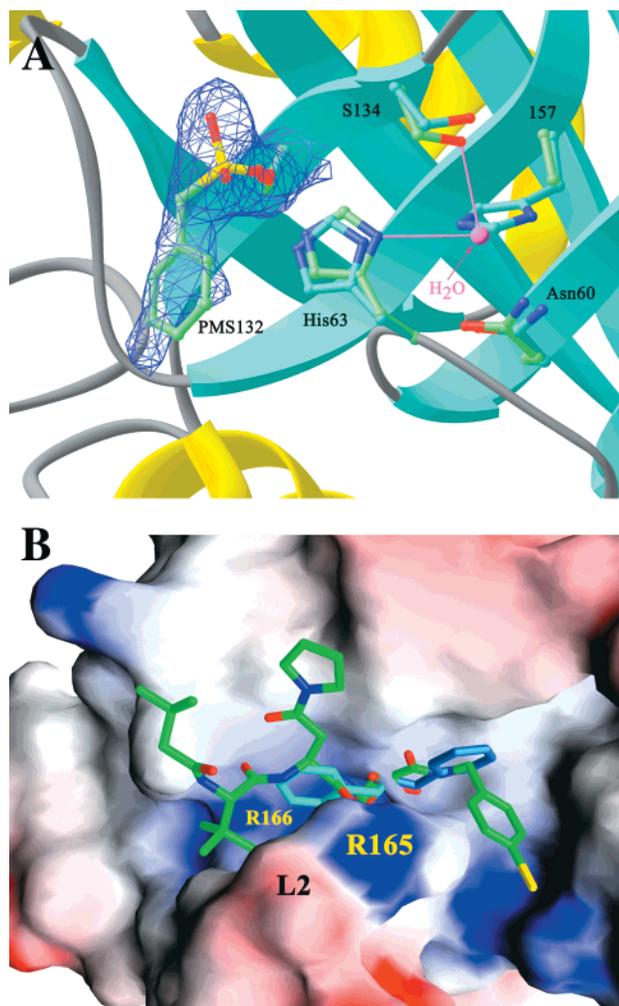


FIGURE 4: Active site of the H157A mutant. (A) Schematic drawing showing the structure of the H157A mutant near the active site. The side chains of residues 132, 63, 157, and 134 are shown for the H157A mutant (in green for carbon atoms) and for the wild-type free enzyme (in cyan). The water molecule that is hydrogen-bonded to the His63 side chain in the H157A mutant is shown as a purple sphere. Also shown is the conformation of the PMSF molecule in the first monomer and its electron density in the final $2F_o - F_c$ electron density map. The contour level is at 1σ . Produced with Ribbons (30). (B) Overlap of the two alternate binding modes of the PMSF molecule (in blue for the first monomer and in cyan for the second monomer) and the peptidomimetic inhibitor (in green). The protease is shown as a molecular surface. Produced with Grasp (31).

Active Site of the H157A Mutant and Binding Mode of PMSF. Of the three mutants, the conformation of the second-member His63 residue in the H157A mutant bears the strongest similarity to the wild-type structure (Figure 4A). The hydrogen bond between the second-member His63 side chain and the third member is eliminated by the mutation. However, the structure shows that a solvent molecule takes up the position of the side chain of the third member. This solvent molecule makes hydrogen bonds to the His63 side chain and the side chain hydroxyl group of Ser134 (Figure 4A).

The crystal structure of the H157A mutant actually contains a covalent adduct with the PMSF inhibitor. The electron density for the sulfonyl portion of this inhibitor is well-defined, and it clearly shows the covalent linkage between PMSF and the catalytic Ser132 side chain (Figure 4A). The electron density of the phenyl ring is however weaker, suggesting possible flexibility in the binding of this portion of the inhibitor. In addition, the phenyl group has different conformations in the two monomers. In the first monomer, it is pointed toward the P_1' direction, whereas in the second monomer it is pointed toward the P_2 direction. Compared to the structure of the peptidomimetic inhibitor complex, the phenyl groups essentially superimpose with the main chain positions of the P_1' and P_2 residues, respectively, in the two monomers (Figure 4B).

Biochemical and Structural Studies of the S134A Mutants. The structure of the H157A mutant reveals the involvement of a solvent molecule that is hydrogen-bonded to the His63 and Ser134 side chains (Figure 4A). It is likely that this solvent molecule is partly rescuing the catalytic activity of the H157A mutant. To assess the functional significance of this water molecule, we have produced the S134A mutant as well as the S134A/H157A double mutant. The Ser134 residue is almost strictly conserved among the herpesvirus proteases, showing variations to Ala only in the herpes simplex virus type 1 (HSV-1) and Epstein–Barr virus (EBV) enzymes. As expected, the S134A/H157A double mutant shows a larger reduction of the catalytic efficiency as compared to the H157A mutant (Table 2), and our structural studies confirmed that the water molecule is no longer present in this double mutant (Table 1). Therefore, our experimental data confirm that the solvent molecule plays a role in the catalysis by the H157A mutant.

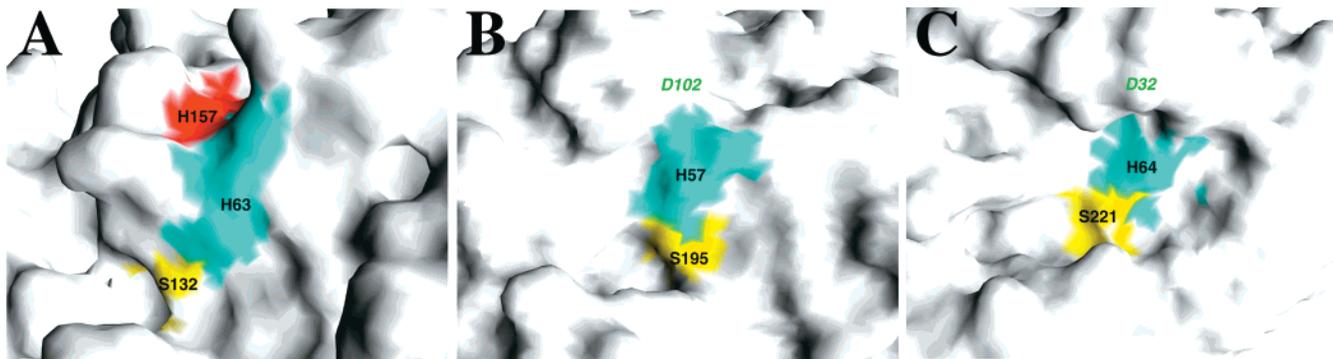


FIGURE 5: The third member of the catalytic triad in classical serine proteases is fully buried. The molecular surfaces near the active site region of (A) HCMV protease, (B) chymotrypsin, and (C) subtilisin are shown colored in yellow for the catalytic serine residue, in cyan for the second-member His residue, and in red for the third member (Asp or His). The structures are aligned such that the catalytic triads are in the same orientation. The lack of a red patch in chymotrypsin and subtilisin shows that the third members are buried in these enzymes. Produced with Grasp (31).

Surprisingly, the S134A mutant itself also shows a significant reduction in the catalytic activity, which is actually more severe than that of the H157A mutant (Table 2). Our kinetic results are consistent with earlier biological experiments showing that the S134A mutation in simian CMV protease can reduce the catalytic activity (27). A slight decrease in the catalytic activity of HCMV protease by the S134A mutation was also observed from transfection experiments (28). To understand the molecular basis for this kinetic observation, we have produced the crystal structure of the S134A mutant of HCMV protease as well (Table 1). The structure shows that the His157 side chain is disordered, and there is no stable hydrogen-bonding partner for the His63 side chain (data not shown). Therefore, the Ser134 side chain may be required to stabilize the conformation of the His157 side chain in HCMV protease. Further studies on the HSV-1 and EBV proteases are needed to understand why they do not depend on a Ser residue at position 134.

A Conserved Solvent Structure Near the Active Site. Analysis of the crystal structures of the His157 mutants, as well as the Ser134 mutants, showed that the positions of three water molecules near the active site are conserved among the structures (Figure 3). Subsequent examination of the free enzyme structures reported earlier showed the presence of the same water molecules, as was also noted in the crystal structure of HSV-2 protease (9). The first of these three water molecules is hydrogen-bonded to the main chain amide of Leu133 and the side chain guanidinium group of Arg166 (W1 in Figure 3). The second water molecule is hydrogen-bonded to the main chain amides of Arg165 (the oxy-anion hole) and Arg166, and the main chain carbonyl of Leu131. The third water molecule is hydrogen-bonded to the Ser132 side chain, and this molecule is replaced by the P₁ residue in the peptidomimetic inhibitor complex (11) and by the sulfonyl oxygen in the PMSF complex with the H157A mutant.

The presence of the W1 and W2 water molecules in the peptidomimetic inhibitor complex could not be confirmed from the structural studies due to the moderate resolution (2.8 Å) of that structure (11). However, simple modeling studies showed that they are likely to be present in the inhibitor complex as well. These two waters are located at the base of the S₁ binding pocket. It might be possible to replace these conserved, structural water molecules with atoms in well-designed inhibitor molecules. This strategy has been used successfully in the design of novel and potent inhibitors that replace a structural water molecule in the active site of HIV-1 protease (29).

CONCLUSION

In an effort to better understand the novel catalytic triad of the herpesvirus proteases, and its implications in the mechanism of serine proteases in general, we have made site-specific mutants of HCMV protease to identify the role of the His157 third member in the triad. Comparisons of the catalytic activities of protease samples that contain functional diads and triads showed that the Ser132-His63 diad of HCMV protease is as functional as the diad in classical serine proteases. However, the triad of HCMV protease is much weaker than that of the classical serine proteases, due to a significantly smaller contribution from the His157 third member.

A weaker hydrogen bond between the His157 and His63 side chains is one of the possible reasons for the diminished role of the third member in HCMV protease. The kinetic results from the H157E mutant suggest, however, that introduction of a strong hydrogen-bonding partner is not enough to increase the activity. A major difference between the catalytic triads in classical serine proteases and HCMV protease is that the side chain of the third member is fully buried in the classical serine proteases (Figure 5). This burial of the acidic third member significantly enhances its interaction with the second-member His residue. In the herpesvirus proteases, however, the third member is exposed to the solvent. The surface exposure of the third member may be another factor that limits the functional importance of the His157 residue in the catalysis by herpesvirus proteases.

The kinetic studies showed that the mutants studied here all have weaker catalytic activity than the wild-type enzyme, suggesting a His residue is preferred by the protease at this position. Removing the His157 third member causes roughly a 150-fold loss in catalytic activity (this excludes the rescuing effect of the water molecule in the H157A mutant), while causing little structural changes in the active site of the enzyme. The H157Q mutant shows subtle structural changes in the active site, but the restoration of a hydrogen bond to the His63 second member leads to about an 8-fold increase in activity. A similar activity was observed in the H157A mutant, where the hydrogen bond is made to an ordered solvent molecule. Introduction of a strong hydrogen-bonding partner in the H157E mutant gives rise to a further 2-fold increase in activity. However, the crystal structure showed changes in the conformation of the His63 and Ser132 residues in the H157E mutant. In addition, the structures of the second monomer of the H157E and H157Q mutants hint at possible flexibility of these side chains, which could also decrease their contribution to the catalysis. It may be possible that His157 is preferred as the third member as it can maintain the best arrangement of the catalytic triad residues in herpesvirus proteases.

ACKNOWLEDGMENT

We thank Yingwu Xu, Gerwald Jogl, and Zhiru Yang for help with data collection at synchrotron radiation sources.

REFERENCES

1. Fields, B. N., Knipe, D. M., and Howley, P. M. (1996) *Fields Virology*, Lippincott-Raven Press, New York.
2. Gibson, W., Welch, A. R., and Hall, M. R. T. (1995) *Perspect. Drug Discovery Des.* 2, 413–426.
3. Waxman, L., and Darke, P. L. (2000) *Antiviral Chem. Chemother.* 11, 1–22.
4. Tong, L., Qian, C., Massariol, M.-J., Bonneau, P. R., Cordingley, M. G., and Lagace, L. (1996) *Nature* 383, 272–275.
5. Qiu, X., Culp, J. S., DiLella, A. G., Hellmig, B., Hoog, S. S., Janson, C. A., Smith, W. W., and Abdel-Meguid, S. S. (1996) *Nature* 383, 275–279.
6. Shieh, H.-S., Kurumbail, R. G., Stevens, A. M., Stegeman, R. A., Sturman, E. J., Pak, J. Y., Wittwer, A. J., Palmier, M. O., Wiegand, R. C., Holwerda, B. C., and Stallings, W. C. (1996) *Nature* 383, 279–282.
7. Chen, P., Tsuge, H., Almassy, R. J., Gribskov, C. L., Katoh, S., Vanderpool, D. L., Margosiak, S. A., Pinko, C., Matthews, D. A., and Kan, C.-C. (1996) *Cell* 86, 835–843.

8. Qiu, X., Janson, C. A., Culp, J. S., Richardson, S. B., Debouck, C., Smith, W. W., and Abdel-Meguid, S. S. (1997) *Proc. Natl. Acad. Sci. U.S.A.* *94*, 2874–2879.
9. Hoog, S. S., Smith, W. W., Qiu, X., Janson, C. A., Hellmig, B., McQueney, M. S., O'Donnell, K., O'Shannessy, D., DiLella, A. G., Debouck, C., and Abdel-Meguid, S. S. (1997) *Biochemistry* *36*, 14023–14029.
10. Reiling, K. K., Pray, T. R., Craik, C. S., and Stroud, R. M. (2000) *Biochemistry* *39*, 12796–12803.
11. Tong, L., Qian, C., Massariol, M.-J., Deziel, R., Yoakim, C., and Lagace, L. (1998) *Nat. Struct. Biol.* *5*, 819–826.
12. Bonneau, P. R., Grant-Maitre, C., Greenwood, D. J., Lagace, L., LaPlante, S. R., Massariol, M.-J., Ogilvie, W. W., O'Meara, J. O., and Kawai, S. H. (1997) *Biochemistry* *36*, 12644–12652.
13. LaPlante, S. R., Bonneau, P. R., Aubry, N., Cameron, D. R., Deziel, R., Grand-Maitre, C., Plouffe, C., Tong, L., and Kawai, S. H. (1999) *J. Am. Chem. Soc.* *121*, 2974–2986.
14. Corey, D. R., and Craik, C. S. (1992) *J. Am. Chem. Soc.* *114*, 1784–1790.
15. Carter, P., and Wells, J. A. (1988) *Nature* *332*, 564–568.
16. Register, R. B., and Shafer, J. A. (1997) *J. Virol.* *71*, 8572–8581.
17. Pinko, C., Margosiak, S. A., Vanderpool, D., Gutowski, J. C., Condon, B., and Kan, C.-C. (1995) *J. Biol. Chem.* *270*, 23634–23640.
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* *276*, 307–326.
19. Tong, L. (1993) *J. Appl. Crystallogr.* *26*, 748–751.
20. CCP4 (1994) *Acta Crystallogr. D50*, 760–763.
21. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110–119.
22. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905–921.
23. Holskin, B. P., Bukhtiyarova, M., Dunn, B. M., Baur, P., DeChastonay, J., and Pennington, M. W. (1995) *Anal. Biochem.* *227*, 148–155.
24. Bonneau, P. R., Plouffe, C., Pelletier, A., Wernic, D., and Poupart, M.-A. (1998) *Anal. Biochem.* *255*, 59–65.
25. Ogilvie, W., Bailey, M., Poupart, M.-A., Abraham, A., Bhavsar, A., Bonneau, P., Bordeleau, J., Bousquet, Y., Chabot, C., Duception, J.-S., Fazal, G., Goulet, S., Grand-Maitre, C., Guse, I., Halmos, T., Lavallee, P., Leach, M., Malenfant, E., O'Meara, J., Plante, R., Plouffe, C., Poirier, M., Soucy, F., Yoakim, C., and Deziel, R. (1997) *J. Med. Chem.* *40*, 4113–4135.
26. Bonneau, P. R., Hasani, F., Plouffe, C., Malenfant, E., LaPlante, S. R., Guse, I., Ogilvie, W. W., Plante, R., Davidson, W. C., Hopkins, J. L., Morelock, M. M., Cordingley, M. G., and Deziel, R. (1999) *J. Am. Chem. Soc.* *121*, 2965–2973.
27. Welch, A. R., McNally, L. M., Hall, M. R. T., and Gibson, W. (1993) *J. Virol.* *67*, 7360–7372.
28. Cox, G. A., Wakulchik, M., Sassmannshausen, L. M., Gibson, W., and Villarreal, E. C. (1995) *J. Virol.* *69*, 4524–4528.
29. Lam, P. Y. S., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bachelier, L. T., Meek, J. L., Otto, M. J., Rayner, M. M., Wong, Y. N., Chang, C.-H., Weber, P. C., Jackson, D. A., Sharpe, T. R., and Erickson-Viitanen, S. (1994) *Science* *263*, 380–384.
30. Carson, M. (1987) *J. Mol. Graphics* *5*, 103–106.
31. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct., Funct., Genet.* *11*, 281–296.

BI010158B